COLUMBIA SPEED DATING STUDY

CAPSTONE ANALYSIS

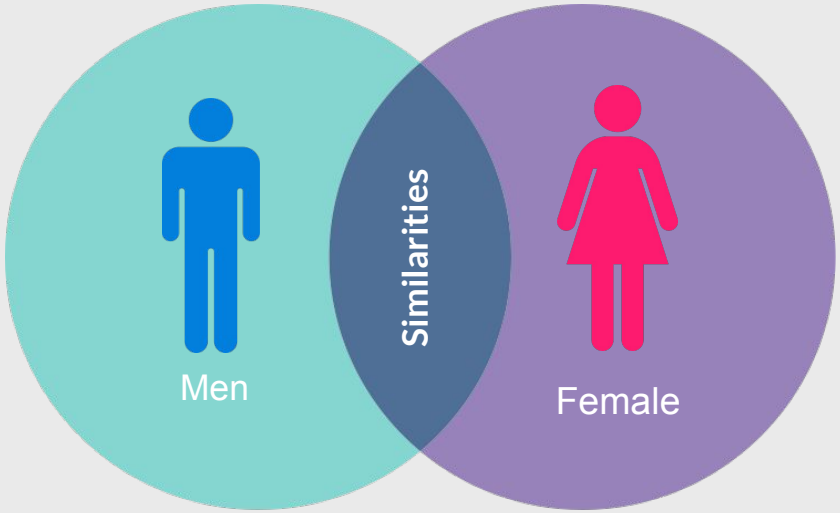
Purpose Of The Case Study?

Business Understanding

Columbia University conducted a research study to monitor and understand Male and Female Behavior when Speed Dating.

Data Representation

Columbia University, 392 single men and women were invited to participate in a speed-dating event. Each participant went on between 10 and 20 four-minute speed dates with members of the opposite sex.



What?

01 What are Male and Female preferences similar when put in a dating environment?

Why?

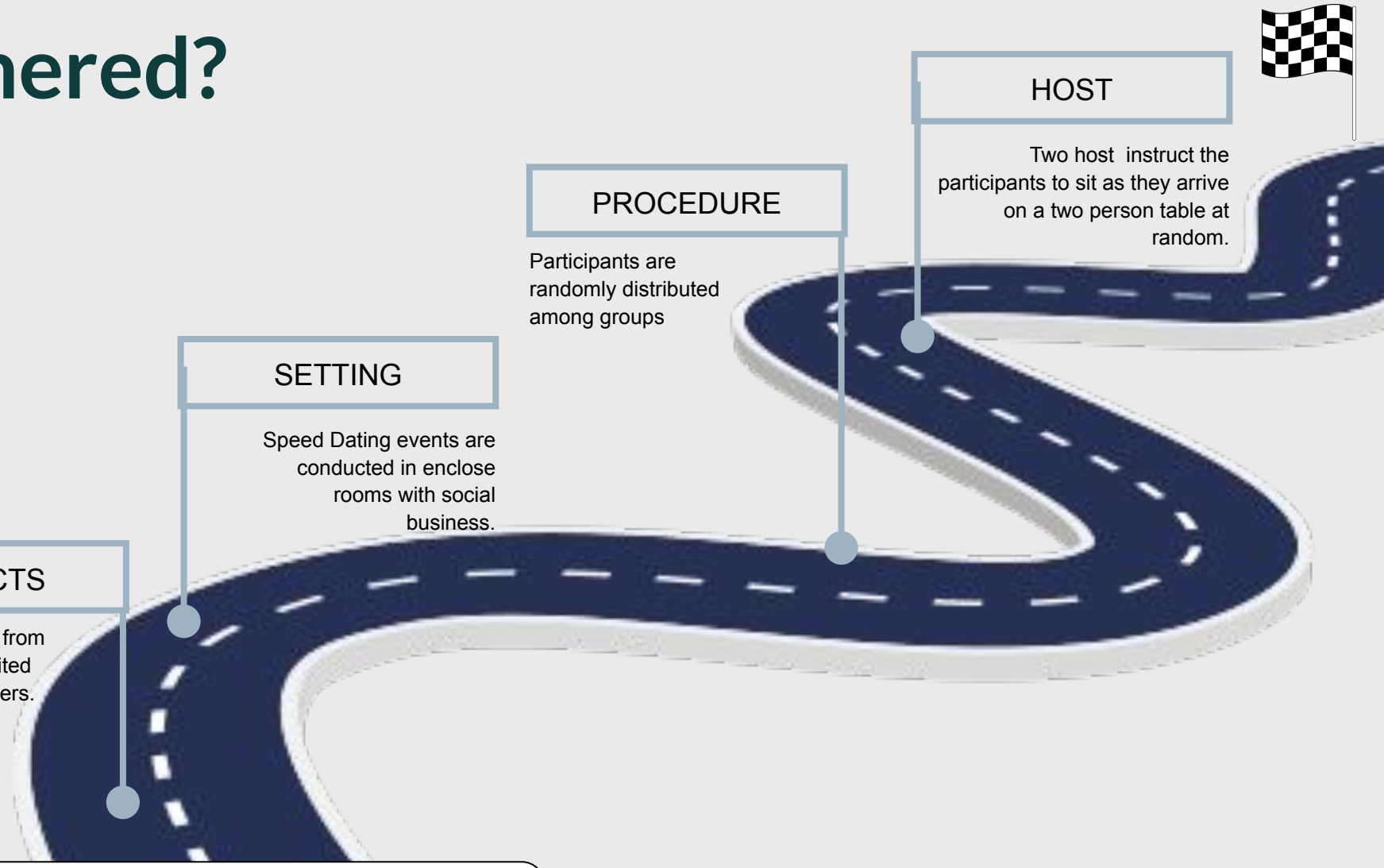
02 Why will this will show us what Men and Women seek in a person regardless of gender.

How?

03 How will this data pair with the regression variable of gender to first understand individual behaviors and then be compared to one another to find correlations.

COLUMBIA

How Was The Data Gathered?



HOST

Two host instruct the participants to sit as they arrive on a two person table at random.

PROCEDURE

Participants are randomly distributed among groups

SETTING

Speed Dating events are conducted in enclose rooms with social business.

SUBJECTS

Subjects are drawn from Columbia or recruited through email or fliers.

COLUMBIA

Case Study Important Details

WAVES	There are 21 waves “sessions” of speed dating	Each wave is held on a seperate date	Days started on Oct 16, 2002	Days ended on April 7th, 2004
PARTICIPANTS	There has been a minimum of 15 participant per session.	There has been a maximum of 32 participants per session.	There is a total of 274 males	There is a total of 276 females
CONDITIONS	Not all waves where controlled, some had interference or conditions.	There are 2 conditions: 1= limited choice 2 = extensive choice	• Limited Choice only provided specific answers.	2. extensive choices allow for a broad amount of answer.

Case Study Limitations

MISSING DATA

Not all the information about the Columbia Speed Dating results are available to us in the CSV file. Because of this we will only research on what we have available, which might be different from original conclusion.

INTERFERENCE

A documents provided to us shows us the days of when waves “session” happen and what variations “interferences” happened, however, we weren’t told if these session should be regarded or disregarded. So we will treat the data as normal, assuming participants still met for 4 minutes and 1 party at least did a review.

Note: Are you wanting to explore this case study

The link for this case study is below of each slide. Because we have to work around limitation our goal again is to understand how the Men and Women behaved by looking at average data of gender wave “session”.

Gender Specific Data

Women in Group (0)

15% of Women 18-42 who speed date are 27 years old and a majority.

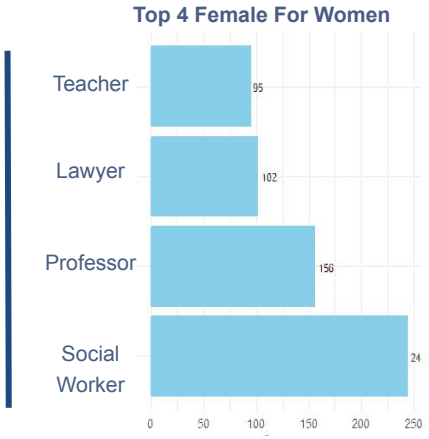
INCOME

$109,031

M $42,225

$8,607

On average Women speed dating have a salary of $42k.



Men in Group (1)

12% of Men 18-55 who speed date are 26 years old and a majority.

INCOME

$106,663

M $45,360

$12,063

On average Men speed dating have a salary of $45k.

Finding Correlations

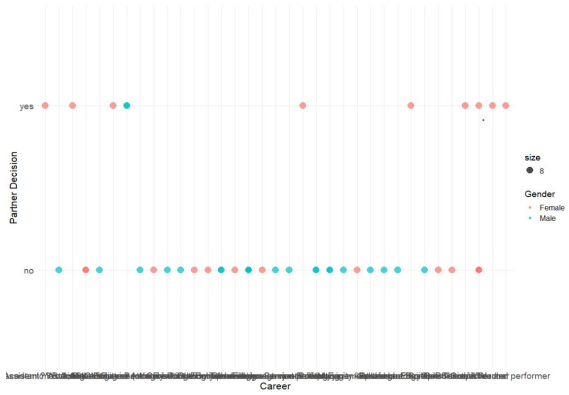
A Partner is likely to say “Yes” or “No” dependent on:

1. Career
2. Jobs
3. Appearance
4. Income

Partner Decision Correlations

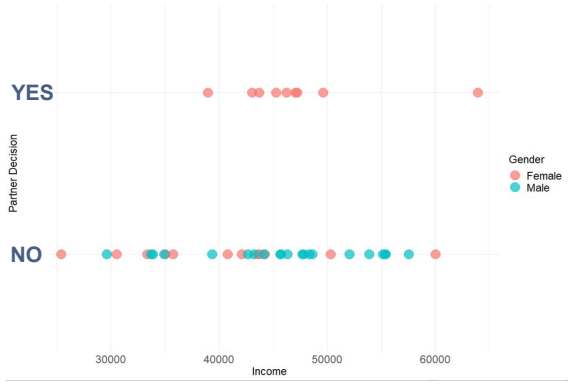
Appearance Rating

A speedsters apperacing played a big role on females passing on a data base on looks.



Career

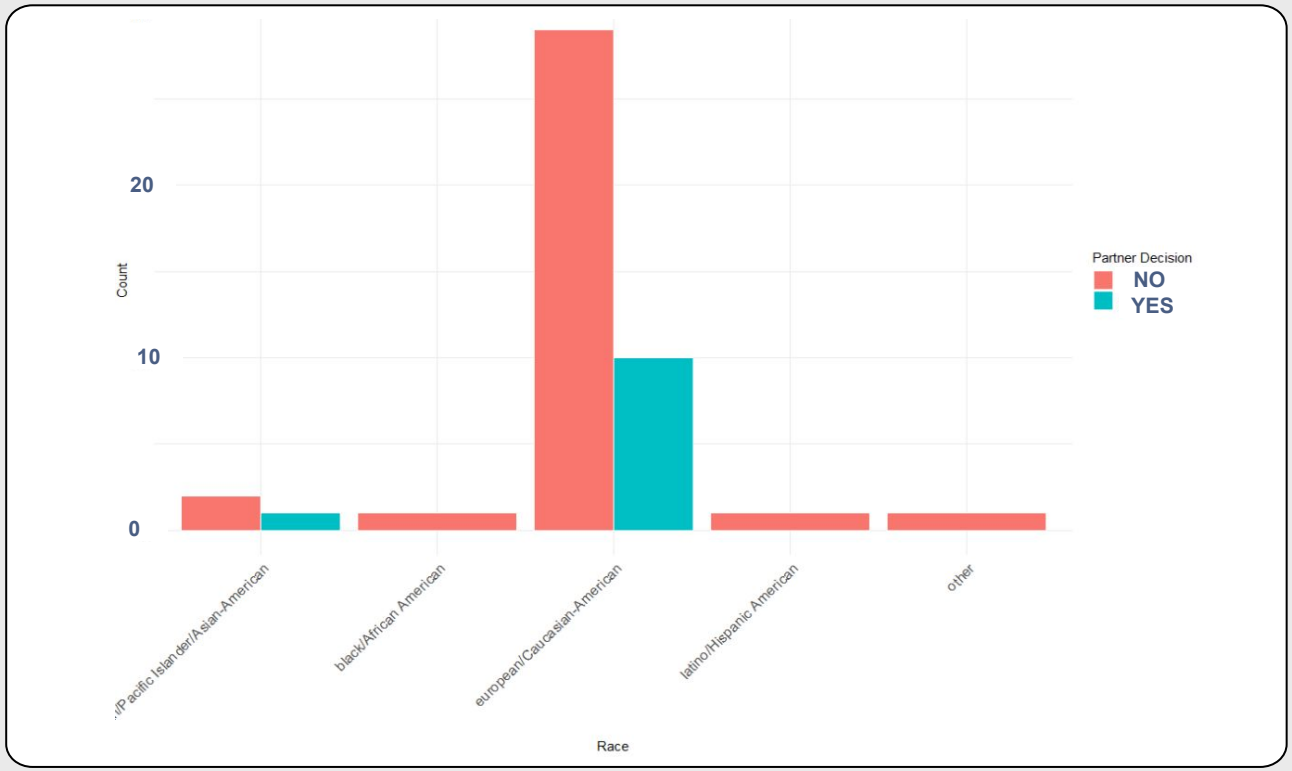
Career has a correlation between males and females that's opposite. Females prefer males with high income jobs, while males do not have a preference.



Income Amount

It is apparent that the female population passed on Male partners when looking at income. Males seem to not leverage income into their decision of passing on a female partner.

Ethnicity is Non-Significant Variable

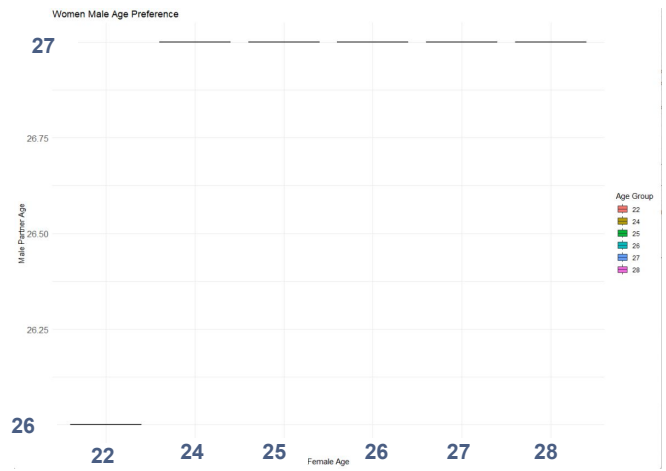


Race & Ethnicity

Race and ethnicity did not really play a big role or significant role on males and females passing on partners. This is apparent when we leverage percentages.

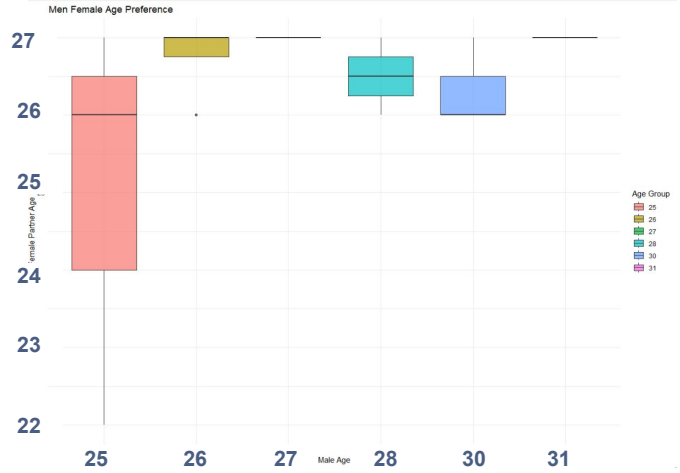
Age is Non-Significant Variable

Women Age Preference



Females ages 23 - 28, find that men 27 year of age are their prefer date. While females age 20-22 prefer man in the age of 26 years of age.

Male Age Preference



Males have the biggest preference on age, with 25 year old men prefer women who are in between the ages of 24-26.5 year of age.

Recommendation: Focus on Ages for decisions

Although the age margin is smaller on women than it is for men, data showed preferences was not correlated by age in Men or Women, meaning it didn't affect their interest score. But nevertheless, understanding the ages that are more popular, helps us understand where the biggest impact with the shortest amount of effort can be measured.

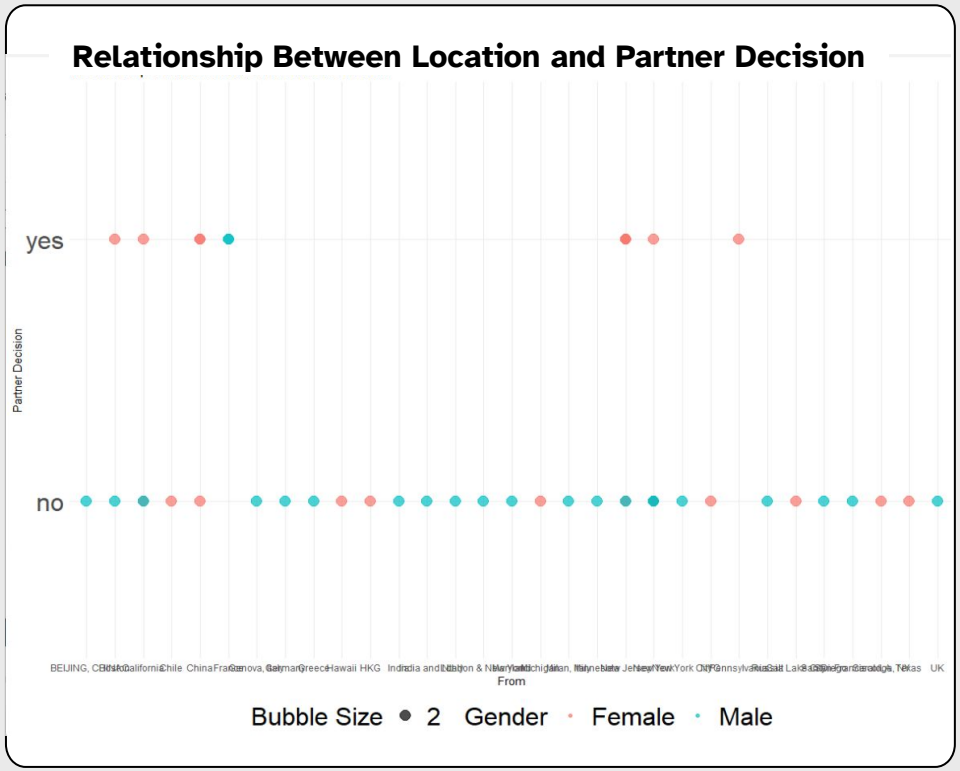
Hobbies Wasn't A Significant Variable

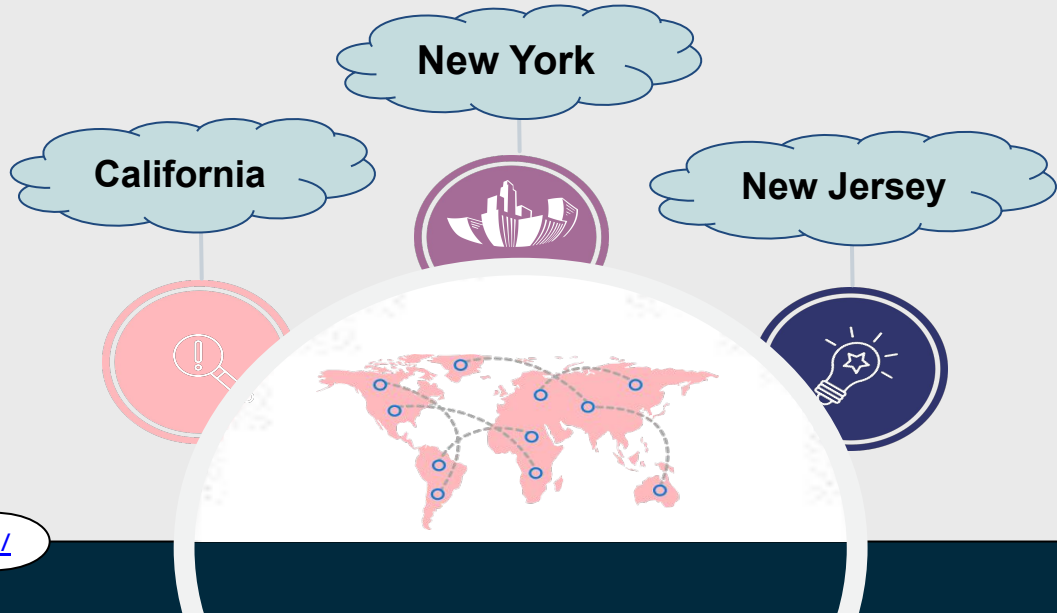
TOP 8 HOBBIES

We see that although people share similar interest on what they like doing, other factor carried more weight on a partners decision like income, career, and appearance.

Location Wasn't A Significant Variable

Here we can see the relationship between locations and if they impacted decision. We can see that here that most of the population did not find the location to be something that is really impactful.





www.stat.columbia.edu/~gelman/arm/examples/speed.dating/

Conclusions

We did analysis and evaluated both Genders in various categories finding the correlation and non-significant relationships. One key finding that was interesting was that Females answers were mostly unpredictable. In almost every calculation, women where on both sides of the spectrum. With more define data, we could have found if female answer would continue bringing in yes after initially meeting with a Male, or if it would discourage, since Females rated to want far more from a Male than the Male expected of the Female.

Correlations

1. Career
2. Jobs
3. Appearance
4. Income

Non-Significant

1. Location
2. Hobbies
3. Age
4. Ethnicity

DATA ACTIONS & PRIORITY

In consequence, this representation ranks the first actions to take in terms of data collection, preparation, and analysis.

1 = First Action → **5 = Last Action**

1	Understand what the final decision purpose for the gathered data is.	Identify
2	Determine which samples in the data will bring confidence in the findings.	Planning
3	Is this a qualitative or quantitative or a mixed-method approach	Data Identifying
4	Time-consuming and error-prone step of the data collection process.	Data Gathering
5	Make decisions based on historical trends and up-to-date information.	Data Analysis